Sex-Dependent Mechanisms For Expansions and Contractions of the CAG Repeat on Affected Huntington Disease Chromosomes

B. Kremer,[1,2] E. Almqvist,[1,3] J. Theilmann,[1] N. Spence,[1] H. Telenius,[1] Y. P. Goldberg,[1] and M. R. Hayden[1]

[1]Department of Medical Genetics, University of British Columbia, Vancouver; [2]Department of Neurology, Academic Hospital, Nijmegen; and [3]Departments of Geriatric Medicine and Molecular Medicine, Karolinska Institute, Huddinge University Hospital, Huddinge, Sweden

Summary

A total of 254 affected parent-child pairs with Huntington disease (HD) and 440 parent-child pairs with CAG size in the normal range were assessed to determine the nature and frequency of intergenerational CAG changes in the HD gene. Intergenerational CAG changes are extremely rare (3/440 [0.68%]) on normal chromosomes. In contrast, on HD chromosomes, changes in CAG size occur in ~70% of meioses on HD chromosomes, with expansions accounting for 73% of these changes. These intergenerational CAG changes make a significant but minor contribution to changes in age at onset ($r^2 = .19$). The size of the CAG repeat influenced larger intergenerational expansions (>7 CAG repeats), but the likelihood of smaller expansions or contractions was not influenced by CAG size. Large expansions (>7 CAG repeats) occur almost exclusively through paternal transmission (0.96%; $P < 10^{-7}$), while offspring of affected mothers are more likely to show no change ($P = .01$) or contractions in CAG size ($P = .002$). This study demonstrates that sex of the transmitting parent is the major determinant for CAG intergenerational changes in the HD gene. Similar paternal sex effects are seen in the evolution of new mutations for HD from intermediate alleles and for large expansions on affected chromosomes. Affected mothers almost never transmit a significantly expanded CAG repeat, despite the fact that many have similar large-sized alleles, compared with affected fathers. The sex-dependent effects of major expansion and contractions of the CAG repeat in the HD gene implicate different effects of gametogenesis, in males versus females, on intergenerational CAG repeat stability.

Introduction

The underlying mutation in almost all persons with HD is an expansion of a CAG trinucleotide repeat beyond

35 repeats within a novel gene on 4p16.3 (The Huntington's Disease Collaborative Group 1993; Kremer et al. 1994). A strong inverse correlation has been shown between CAG repeat size and onset of HD, with a more expanded repeat associated with earlier age at onset (Andrew et al. 1993; Duyao et al. 1993; Snell et al. 1993; Stine et al. 1993; Zühlke et al. 1993). This correlation is most significant for persons with onset before the age of 20 years (juvenile HD) but does not have any significant association for persons with onset after age 60 years (Kremer et al. 1993; Telenius et al. 1993).

Different studies have shown that the expanded CAG trinucleotide is unstable during transmission through the germ line (Andrew et al. 1993; Duyao et al. 1993; Snell et al. 1993; Stine et al. 1993; Zühlke et al. 1993). This is particularly evident for persons with juvenile HD, where the sex of the transmitting parent has a major influence on CAG expansion (Telenius et al. 1993). The molecular basis for this previously described phenomenon of predominance of paternal descent for juvenile HD was shown to be due to significant amplification of the repeat during transmission through the male germ line. However, small intergenerational changes in CAG repeat length are more common.

This study was designed to accurately assess the frequency of intergenerational CAG changes in a large cohort of HD and control parent-child pairs of different ages. We have attempted to determine those factors which might influence the stability of CAG repeat size when transmitted from parent to child. In an effort to assess these factors as accurately as possible, DNA from 254 affected parents and offspring with CAG sizes in the HD range were analyzed on the same gel, in an effort to discern small changes in CAG repeat size. Here we show on HD chromosomes that intergenerational changes occur in ~70% of meioses, with expansions accounting for 73% of all of these changes. In contrast, on normal chromosomes with CAG size ranges of 10–28 CAG repeats, intergenerational CAG changes are extremely rare (3/440 [0.68%]). Furthermore, while changes in CAG size on HD chromosomes are associated with changes in age at onset, this effect is minor, with intergenerational CAG changes accounting for only 19% of the changes

in age at onset. The size of the CAG repeat in the affected parent was associated with larger intergenerational expansions (>7 CAG repeats), but CAG size made no contribution to the likelihood of smaller expansions (≤7 CAG repeats) or contractions. Smaller expansions occur relatively more frequently in offspring of affected males than in offspring of affected females, while large expansions (>7 CAG repeats) occur almost exclusively in paternal offspring (25/26 [96.1%]). In contrast, contractions occur more frequently in offspring of affected mothers. These findings suggest that mechanisms for intergenerational CAG changes in HD are strongly influenced by the sex of the transmitting parent.

Methods

Selection of Patients and Controls

DNA samples from persons with the diagnosis of HD have been collected from families of different descent, including persons of western European, Black, Chinese, Japanese, and Finnish origins. The diagnosis of HD was made by a neurologist or geneticist. Control chromosomes were from unrelated persons of similar descent, including spouses of affected persons and unrelated individuals.

Of 1,498 affected persons in our database, a total of 254 parent-child pairs were identified in which DNA was available from both affected parent and his or her offspring and where both had CAG sizes in the affected range (>35 repeats). The cohort came from 128 unrelated families. A total of 56 nuclear families in whom an affected parent and two or more children had CAG sizes in the HD range were included. In addition, a total of 440 parent-child pairs with CAG sizes in the normal range (10–28) were studied. Eleven parent-child pairs were excluded by virtue of the fact that nonpaternity or sample mix-up confounded these analyses.

DNA Analysis and Assessment of CAG and CCG Repeat Lengths

DNA was extracted from leukocytes by standard procedures (Kunkel et al. 1977). In the past, CAG trinucleotide assessment was performed by using primers encompassing both CAG (Andrew et al. 1993; Goldberg et al. 1993a) and the flanking CCG repeat (Andrew et al. 1994). However, in this study, a precise assessment of the number of CAG repeats was performed by excluding the CCG repeats (Squitieri et al. 1994). Intergenerational changes in repeat size were confirmed by analysis of parent-child pairs on the same polyacrylamide gel.

To directly assess only the CAG repeat size in the HD gene, primer pair HD344/HD450 immediately flanking the CAG repeat was used. Primer sequences were HD344 (5' CCTTCGAGTCCCTCAAGTCCTTC 3') and HD450 (5' GGCGGCGGTGGCGGCTGTTG 3').

PCR was performed in a 25-μl volume, with 100 ng of genomic DNA in 2 mM $MgCl_2$, 50 mM KCl, 20 mM Tris pH 8.4, 17% glycerol, 1% formamide, 150 μM dATP, 150 μM dCTP, 150 μM dGTP, 150 μM dTTP, 20 pmol of both primers, 0.5 pmol of $\gamma^{32}P$ end-labeled HD344, and 5 U of *Taq* polymerase. Cycling parameters were 95°C for 3 min, followed by 35 cycles at 94°C for 1 min, 64°C for 1 min, and 72°C for 1 min, with a final extension at 72°C for 5 min. PCR products were sized by using an M13 sequencing ladder as well as cloned PCR products of known-size CAG sequences.

Statistical Analysis

The sizes of the CAG and CCG alleles were entered into a database containing clinical information and were subjected to further analysis. To examine association between various variables, simple linear regression was used. Comparisons of proportions were performed by χ^2 analysis. To examine the influence of sex of affected parent and behavior of the CAG repeat during transmission (the categories were "expansion over 7," "expansion between 1 and 7," "no change," and "contraction"), analysis of variance with an incomplete factorial design was used, since one cell contained one case only. Post hoc comparisons were performed according to the procedure of Tukey (1977). In the other procedures no adjustment was made for multiple comparisons; the α-level was set at .05.

Results

CAG Sizes in Affected Parents and Offspring

DNA was available from 119 affected father-child pairs and 135 affected mother-child pairs, for a total of 254 parent-child pairs (table 1). No data were inferred. The mean onset age in the parents was higher (42.8 years; range 18–75 years; $n = 113$) than in the offspring (31.1 years; range 4–54 years; $n = 81$) (table 1). This might be explained in part by ascertainment bias, since those offspring in the sibship destined to have later onset were not included. Although the mean CAG size in the 168 parents was smaller (43.4 repeats) than in the offspring (46.3 repeats) (table 1), the median size was identical in both groups, i.e., 43 repeats. This can be explained by the skewed nature of the CAG size distribution in both groups, which was similar in both groups except for a number of very long repeats in the offspring. The maximum parental upper allele size was 58 CAG repeats. Thus, except for a number of offspring with >58 CAG repeats, the distribution of parental and children's CAG sizes looked similar.

Intergenerational Instability of CAG Repeat

Although the median CAG length did not change between the parent group and the offspring group, CAG

Table I

Demographics of Cohort

	No.
Affected parent:	
Father	119
Mother	135
Total affected parent-child pairs	254
Parent-child pairs with CAG size in normal range	440
Separate families	128
CAG size (168 parents [79 female and 89 male]):	
Mean	43.4
Median	43
Range	37–58
254 Offspring:	
Mean	46.3
Median	43
Range	37–121

	Age at Onset (years)
Parents (n = 113):	
Mean	42.8
Median	40
Range	18–75
Offspring (n = 81):	
Mean	31.1
Median	33
Range	4–54

repeat lengths studied in individual parent-child transmissions were often unstable (70.4%) (table 2). In 46 instances (18.1%) the CAG length in the offspring was decreased, in 75 instances (29.5%) CAG length remained constant, and in 133 instances (52.3%) CAG length increased (table 2). Decreases in triplet-repeat size ranged from −1 to −4, whereas increases ranged from +1 to +74 (fig. 1). The vast majority (90%) of transmissions had CAG changes ranging from −4 (contraction) to +7 (expansion). A total of 10.2% (26/254) had CAG expansion of >7 CAG repeats (table 2). The range was 8–74 repeats, and 16 of these were associated with onset at or before age 25 years, whereas in the remaining 10 cases the onset age was unknown. Age at onset was not available for 173 persons, either because symptoms had not yet been manifested or because information on age at onset was unavailable.

Influence of Sex of Affected Parent on Intergenerational Change in CAG Size

The sex of the affected parent was a major factor influencing CAG instability. There was a significantly greater likelihood of the CAG repeat length to show an increase in size when it was transmitted from the father (81/119 [68.1%]), compared with when it was transmitted from the mother (52/135 [38.5%]) ($P < 10^{-6}$) (table 2 and fig. 2). This significant difference was primarily due to large expansions (>7 CAG repeats), where 25/119 (21.0%) of paternal transmissions showed such expansion, compared with only 1/135 (0.7%) offspring (who had an increase of 16 CAG repeats) from an affected mother ($P < 10^{-4}$) (table 2 and fig. 3). Analysis of those CAG changes of 1–7 CAG repeats still revealed a trend to an increase in the frequency of paternal descent (56/119 [47.1%] compared with 51/135 [37.8%]), but this did not reach statistical significance ($P = .13$) (table 2 and fig. 2).

When the CAG repeat was transmitted from an affected mother to the offspring, there was a significantly greater chance for this to be stable (49/135 [36.3%]), compared with when the father transmitted the CAG repeat to the offspring (26/119 [21.9%]) ($P = .012$) (table 2). Furthermore, mothers were more likely than fathers to pass on a CAG repeat that was reduced in size (34/135 [25.2%] vs. 12/119 [10.1%]; $P = .002$).

Our database may have some biases of ascertainment, in that we have specifically collected affected parent-child pairs where the offspring had juvenile onset (i.e., before the age of 20 years). A total of 18 such pairs were in this data set. Analysis of the data without these pairs revealed similar statistically significant results, with offspring of affected fathers more likely to have major expansions (>7 CAG repeats) and with offspring of affected mothers more likely to have either no change or contractions. However, in contrast to the results for the analysis of the whole group (table 2), in this group the offspring of affected fathers were also more likely to demonstrate small CAG expansions (≤7 CAG repeats) ($\chi^2 = 6.49$; df = 1; $P = .01$) than were offspring of affected mothers.

Influence of Parent CAG Repeat Size on Stability of CAG Repeat in Offspring

A highly significant correlation between CAG instability and the size of the parental CAG repeat was found for paternal transmissions ($n = 119$; $r = .49$; $P < 10^{-7}$) but not for the maternal transmissions ($n = 135$; $r = .08$; $P = .37$). However, the paternal CAG size did not show a normal distribution, and the significant correlation was influenced by a small number of large expansions from fathers with large CAG sizes (fig. 3).

We next sought to determine whether parental CAG size significantly influenced the likelihood of a particular CAG change in the offspring. Analysis was performed by grouping cases according to sex of affected parent and intergenerational behavior of the CAG (expansion >7 CAG repeats, expansion ≤7 CAG repeats, no change, or contraction) (table 2). Analysis of variance showed a significant difference between the mean paren-

Table 2

CAG Sizes and Sex-of-Parent Effect on Intergenerational CAG Changes
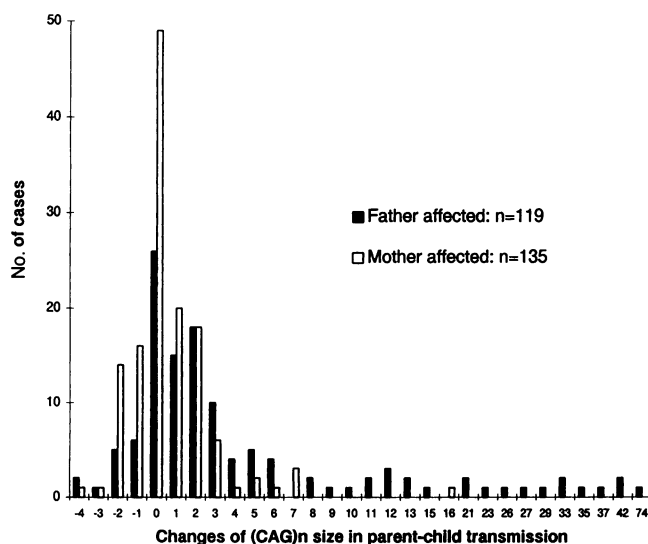
	AFFECTED PARENT					
CAG SIZE	Father	No. (%) of Affected Families	Mother	No. (%) of Affected Families	TOTAL	No. (%) of Affected Families
Expansions >7 repeats:						
Mean ± SD	47.1 ± 4.7 ⎫	25 (21.0)	44	1 (.7)	47.0 ± 4.7	26 (10.2)
Range	39–58 ⎭					
Expansions ⩽7 repeats:						
Mean ± SD	42.2 ± 2.7 ⎫	56 (47.1)	43.7 ± 3.8 ⎫	51 (37.8)	42.9 ± 3.3	107 (42.1)
Range	37–48 ⎭		39–56 ⎭			
No change:						
Mean ± SD	42.3 ± 3.5 ⎫	26 (21.8)	42.7 ± 3.1 ⎫	49 (36.3)	42.5 ± 3.2	75 (29.5)
Range	39–43 ⎭		34–56 ⎭			
Contractions						
Mean ± SD	42.0 ± 1.8 ⎫	12 (10.1)	43.7 ± 2.3 ⎫	34 (25.2)	43.3 ± 2.3	46 (18.1)
Range	39–44 ⎭		39–47 ⎭			
Overall	43.2 ± 2.9	119 (100)	43.5 ± 5.2	135 (100)		254 (100)

NOTE.—Post hoc multiple comparisons between the different cells showed that only fathers with CAG expansions >7 repeats have a significantly different CAG size ($P < .005$) compared with those for any other group. Analysis of variance for mean parental CAG, using sex of affected parent and behavior of intergenerational CAG size as grouping variables, gave the following results: for affected parent—F-ratio = 0.038; $P = .861$; and for intergenerational CAG change—F-ratio = 12.77; $P = 10^{-7}$.

tal CAG size, among all four groups (i.e, expansion >7 CAG repeats and expansion ⩽7 CAG repeats), contractions, and no change in CAG allele size) (F-ratio = 12.77; $P = 10^{-7}$). However, when the two groups of parents, i.e., fathers versus mothers, were compared, there was no difference in the mean CAG size (F-ratio

= 0.038; $P = .861$) (table 2). Post hoc analysis revealed that the group of fathers whose alleles expanded >7 CAG repeats in their offspring had a significantly larger CAG repeat length than did any other group (in all cases $P < .005$) (table 2). In contrast, none of these other groups showed any differences among themselves.

From these analyses, it is apparent that fathers—but not mothers—with relatively large alleles may transmit further expanded alleles to their offspring. Therefore, a

Figure I Distribution of intergenerational changes of CAG size in 254 affected parent-child transmissions. A bell-shaped curve with an abnormally long tail comprising large expansions is evident.

Change in allele size on transmision by sex of affected parent
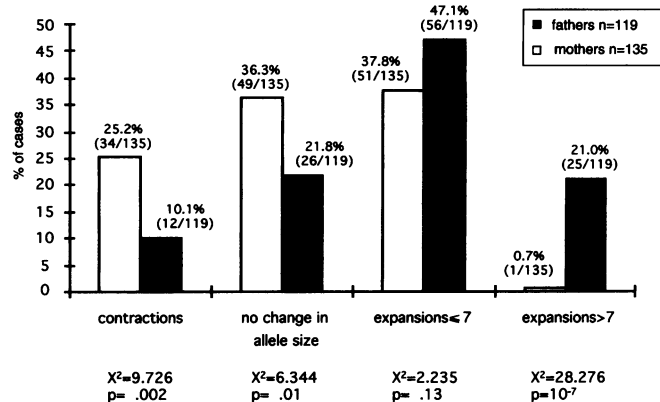
Figure 2 Analysis of the influence of sex of origin on intergenerational CAG changes in 254 parent-child pairs
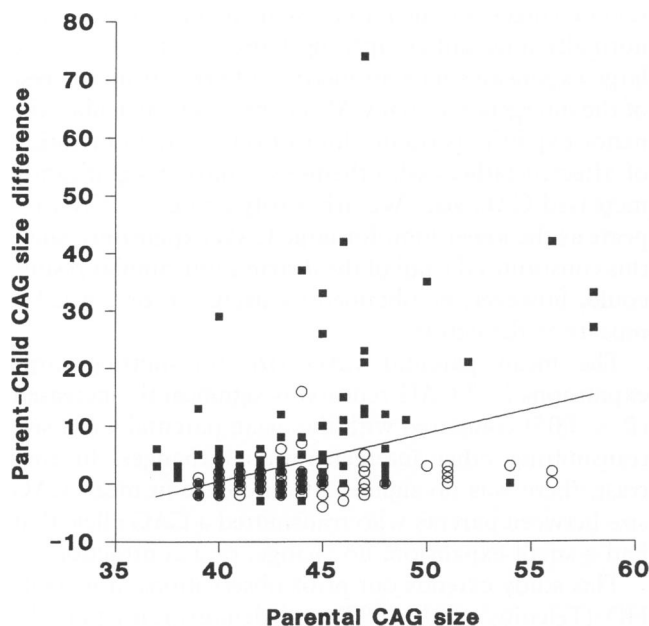
Figure 3 Relationship between parental CAG size and intergenerational change in CAG. A highly significant correlation between CAG instability and the size of the CAG repeat was found for paternal transmissions ($r = .049; P > 10^{-7}$) but not for maternal transmissions ($r = .08; P = .37$). Paternally derived CAG alleles, at every CAG size, are more likely to have a greater expansion than are maternally derived CAG alleles, highlighting the critical role of sex of origin as the primary determinant of parent-child CAG size difference.

larger parental CAG size itself is not sufficient for these primary expansions, since these occur only in transmission from fathers. For example, 22 mothers had CAG size >45, and none of their offspring had a large CAG expansion (fig. 3). The importance of sex of origin is further illustrated in the case of the small expansions or contractions. No effect of CAG size but, again, a clear effect of parental sex can be demonstrated (fig. 3).

Influence of Sex of Affected Grandparent on Intergenerational Instability

The sex of the affected grandparent was known in 108 of the 254 parent-child pairs, including 56 grandmothers and 52 grandfathers. The sex of the grandparent had no significant effect on meiotic instability of the CAG repeat. There was a slight preponderance of affected grandmothers for offspring with CAG expansion $\leqslant 7$ CAG repeats, and there were more affected grandfathers for offspring with contractions of CAG repeats, but these differences did not reach statistical significance (table 3). The sex of the offspring clearly had no affect on intergenerational CAG instability ($\chi^2 = 5.51$, df = 2, $P = .138$).

Influence of CAG Repeat Size of Unaffected Chromosome on Intergenerational Instability

Correlational analysis revealed that neither the CAG size on the normal chromosome of the affected parent

($r = .012; P = .85$) nor the CAG size of the normal chromosome of the offspring ($r = .12; P = .06$) had any influence on CAG instability of the affected chromosome.

Association between Intergenerational Change in CAG Size and Change in Age at Onset

An important question is whether intergenerational CAG change is associated with changes in age at onset of disease. For this analysis, 65 parent-child pairs were available in whom both the parent's and the child's age at onset were known. A significant association existed between the difference in onset age and the intergenerational difference in CAG repeat ($n = 65; r = .43; r^2 = .19; P = .0004$) (fig. 4). This was apparent for both affected mothers ($n = 33; r = .42; r^2 = .17; P = .01$) and affected fathers, although in the latter it barely reached significance ($n = 32; r = .35; r^2 = .12; P = .05$). The same association was demonstrated ($n = 58; r = .33; r^2 = .11; P = .01$) after expansions >7 CAG repeats were omitted from the analysis (since they could constitute outliers). Thus, intergenerational changes in CAG size do contribute to anticipation but are not the major influences of intergenerational changes in ages at onset.

The confidence limits around these predicted mean CAG changes are extremely broad (fig. 4). For example, although in these 65 cases the mean difference in CAG allele size was 5.3 repeats and the mean difference between onset age in affected offspring and that in parents was 10.2 years, the 95% confidence limit around this mean is 17.0 years. Thus, clinically it is impossible to accurately predict change in age at onset, on the basis of the difference in CAG size between affected parents and offspring.

### Intergenerational Changes of CAG Repeat on Normal Chromosomes

A total of 450 parent-child pairs with a normal size CAG repeat were assessed. The range of CAG repeats

Table 3

Influence of Sex of Affected Grandparent on Intergenerational CAG Instability

CAG SIZE IN AFFECTED OFFSPRING	NO. (%) OF CASES WITH AFFECTED GRANDPARENT		
	Grandfather	Grandmother	TOTAL
Expansion >7	9 (17.3)	7 (12.5)	16 (14.8)
Expansion ≤7	22 (42.3)	36 (62.5)	57 (52.8)
No change	13 (25.0)	11 (19.6)	24 (22.2)
Contraction	8 (15.4)	3 (5.4)	11 (10.2)
Total	52 (100)	56 (100)	108 (100)

NOTE.—$\chi^2 = 5.514$; df = 2; $P = .138$.
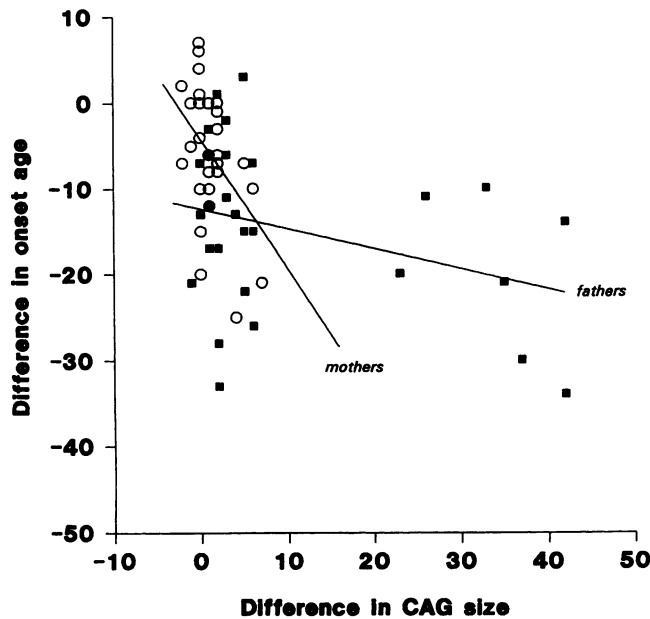
Figure 4   Association between intergenerational change in CAG size and change in age at onset. This association is particularly evident for offspring of affected mothers ($r^2 = .17$; $P = .01$) and just reached significance for offspring of affected fathers ($r^2 = .12$; $P = .05$). This reveals that intergenerational changes in CAG size do contribute to the anticipation but are not the major influences of intergenerational changes in age at onset.

for normal chromosomes was 10–28. Initially there were a total of 81 meioses that revealed potential intergenerational CAG changes. All of these DNAs were analyzed again on the same gel next to each other, to facilitate detection of any changes. This analysis reduced the number of potential CAG changes to 13 parent-child pairs. These pairs were assessed by using a highly polymorphic CA repeat from chromosome 4p (data not shown), and this revealed that, in 10 instances, either nonpaternity or sample mix-up accounted for the intergenerational difference in CAG size. However, in three parent-child pairs a CAG change was detected. These included two expansions of one CAG triplet from 19 to 20 on one chromosome from 16 to 17 and from 25 to 28 on another. In all of these instances, the transmitting parent was the father.

Discussion

We have shown that intergenerational CAG changes in the HD gene occur in ~70% of meioses. In 52.1% of cases (133/254) these changes are expansions, whereas in 18.1% (46/254) a decrease in CAG size has been documented. In contrast, only three CAG intergenerational changes were seen in 440 normal parent-child pairs (0.68%).

Assessment of the distribution of intergenerational

CAG changes revealed a bell-shaped curve with an abnormally long tail comprising large expansions. These large expansions have properties different from the rest of the intergenerational CAG changes. In particular, the major expansions occur almost exclusively in offspring of affected fathers who themselves have a significantly increased CAG size. We arbitrarily chose >7 CAG repeats as the lower limit for large CAG expansions, since this constitutes the tail of the distribution. Similar results could, however, be obtained by using >6 or >8 CAG repeats as the cutoff.

The mean parental CAG size transmitting large expansions (>7 CAG repeats) is significantly increased ($P < .005$) compared with the mean parental CAG size transmitting other intergenerational changes. In contrast, there was no significant difference in mean CAG size between parents who transmitted a CAG allele that had a small expansion, no change, or a contraction.

This study extends our prior observations in juvenile HD (Telenius et al. 1993) and demonstrates that the sex of the parent is the major determinant for all CAG intergenerational changes in the HD gene. Offspring of affected fathers are much more likely than are offspring of affected mothers to have large expansions. In contrast, offspring of affected mothers are more likely than are offspring of affected fathers to have either (a) no change or (b) a reduction in size of their CAG repeat.

There have been prior, smaller studies of CAG intergenerational stability in the HD gene. Noveletto et al. (1994) studied 47 affected parent-child pairs, and Zühlke et al. (1993) studied a total of 54 meioses. The findings of our study are similar to the results of these smaller studies, but they are different than those in a study of 55 parent-child pairs (Trottier et al. 1994) that revealed instability in only 40% of the sample.

We have also shown here that intergenerational CAG change is associated with changes in age at onset. This was significant for female transmission ($P = .01$) but was only of marginal significance for male transmission ($P = .05$). However, intergenerational CAG change is associated with only 19% of the change in age at onset and therefore should not at this stage be regarded as the major determinant of anticipation in this group. Whereas, for the group, CAG change is associated with age at onset, this, however, does not have any predictive value, since the confidence limit around these changes are extremely broad.

An important question is whether this sample of parent-child pairs from HD families is representative of the whole HD population. In this study, both affected parents and affected children served as probands. Although attempts were made to obtain blood samples from all offspring of each affected parent, it was difficult to obtain blood from unaffected offspring. Thus, cases destined to have late onset in each sibship are underrepresented in

this cohort. Similarly, an affected person with late onset may not have had a living, affected parent available. This may explain the lack of offspring with an onset age later than 10 years after the onset age in the affected parent. In addition, it is noteworthy that the majority of offspring had onset earlier than that in the affected parent (fig. 3). This bias in ascertainment could have influenced the data in favor of a significant association between age at onset and CAG change; and, until all members of the sibship have been included in the analysis, we would regard the association between intergenerational CAG change and age at onset as not conclusive.

The sex-dependent effects of major expansions and contractions suggest different mechanisms for expansion and contraction of the CAG repeat in the HD gene. We have recently proposed a multistep model for the evolution of the HD chromosome, with multiple small incremental increases in CAG repeat size eventually resulting in an HD allele of intermediate size, which then undergoes massive expansion into the HD range (Squitieri et al. 1994; Almqvist et al. 1995). Interestingly, sex dependence of large expansion of these intermediate alleles (IA) is also apparent (Goldberg et al. 1993*a*). Analysis of 16 new mutations has revealed that, in all instances, the molecular basis for new mutations is an IA that expands significantly into the affected range occurring exclusively during paternal transmission (Goldberg et al. 1993*b*; De Rooij et al. 1993; Myers et al. 1993; Zühlke et al. 1993; authors' unpublished data). Similarly, in the affected range, large expansions of the CAG repeat length occur almost exclusively during paternal transmission. This paternal sex effect might therefore indicate a similar mechanism for both the evolution of a new mutation for HD and HD resulting from large expansion of a CAG allele in the affected range.

We have previously provided evidence for mitotic instability of the CAG repeat in the HD gene (Telenius et al. 1994). This was most predominant in the areas of the brain as well as sperm, where the instability was greater than that seen in the testis. We have also proposed that spermatogenesis itself may be important to the CAG instability in sperm and may provide some insights into the sex-dependent effects of major CAG expansion (Telenius et al. 1994). These include intergenerational changes of IAs, which expand into the affected range, as well as CAG repeats already in the affected range, in cases where the mutant allele was inherited from the father. The fact that offspring of affected mothers almost never receive a significantly expanded CAG repeat, despite the fact that they may have similar sized CAG alleles, implicates different effects of gametogenesis, in males versus females, on intergenerational CAG repeat stability.

Sperm cells continue to have multiple divisions throughout adult life, whereas oocytes are formed by the time of birth and do not undergo additional divisions. CAG mosaicism in oocytes has not yet been assessed for expanded alleles, because of unavailability of such tissue, but would be predicted to be low. If somatic mosaicism was in part a reflection of continued cell division, this could then explain the higher numbers of CAG repeats and greater CAG instability in sperm. Offspring of such males are at risk of inheriting a significantly expanded allele from their father (Telenius et al. 1995). However, the greater mosaicism in sperm compared with other tissues—such as blood, liver, and bowel— in which there is also significant cell turnover would suggest that cell division alone is not the only factor promoting mitotic instability.

There is no evidence that males with HD who have expanded CAG repeat sizes have a limited capacity for fertilization. Another potential reason for the absence of offspring from females with major expansions is that, in some way, massive expansion of the CAG may be detrimental specifically to the oocyte and therefore results in some impairment of fertilization. Whereas the capacity for CAG instability may in part be dependent on cell division, in all likelihood other tissue-specific factors are likely to play a role in determining both the extent of somatic mosaicism in sperm and, consequently, the presence of major intergenerational expansions in offspring of affected males.

In contrast to expansions, contractions are more likely to occur during female transmissions. The significantly different sex-of-origin effects for expansions compared to contractions suggests different mechanisms for expansions versus contractions. We have recently investigated whether deficiencies of the mismatch-repair system could result in patterns of stability similar to either of these changes.

We found that tumor DNA from patients with a deficiency in the mismatch-repair system have frequent CAG changes at the HD locus (80% [8/10]), compared with blood DNA. Six of these eight changes represented small contractions ($\leqslant 3$ CAG repeats) in tumor DNA compared with blood DNA (M. R. Hayden, Y. P. Goldberg, and B. Vogelstein, unpublished data). Furthermore, contractions were also more frequent in other trinucleotides, suggesting that defects in mismatch repair might generally underlie small changes in CAG length, particularly contractions. In these instances CAG size itself plays some role in promoting instability. For example, in the series analyzed it was the longer CAG repeat lengths that were most likely to be unstable. This would suggest that, as the size of the trinucleotide repeat increases, this would be more likely to lead to DNA polymerase slippage during replication. Furthermore, it is noteworthy that large CAG expansions or contractions were not observed in these mismatch-deficient tumors. These data suggest that mismatch defects are more likely

to result in small (but not large) changes in CAG length, particularly contractions. However, sex-of-origin effects on DNA repair are not yet established but could be implicated as a reason for the predominance of maternal descent for contractions.

In summary, the findings in this study conclusively demonstrate that the sex of the affected parent is the major determinant of intergenerational CAG change in the HD gene. Offspring of affected fathers are more likely to have expansions. By contrast, offspring of affected mothers are more likely to have either no intergenerational change or a reduction in CAG size. CAG size itself is a secondary factor, with larger CAG repeat lengths being associated with major expansions and with smaller CAG lengths more likely to demonstrate no intergenerational change.

Acknowledgments

References

Almqvist E, Spence N, Nichol K, Andrew SE, Vesa J, Peltonen L, Anvret M, et al (1995) Ancestral differences in the distribution of the Δ2642 glutamic acid polymorphism is associated with varying CAG repeat length on normal chromosomes: insights into the genetic evolution of Huntington disease. Hum Mol Genet 4:207–214

Andrew SE, Goldberg YP, Kremer B, Telenius H, Theilmann J, Adam S, Starr E, et al (1993) The relationship between trinucleotide repeat length (CAG) and clinical features of Huntington disease. Nat Genet 4:398–403

Andrew SE, Goldberg YP, Theilmann J, Zeisler J, Hayden MR (1994) A CCG repeat polymorphism adjacent to the CAG repeat in the Huntington disease gene: implications for diagnostic accuracy and predictive testing. Hum Mol Genet 3:65–67

De Rooij KE, De Koning Gans PAM, Skraastad MI, Belfroid RDM, Vegter-Van Der Vlis M, Roos RAC, Bakker E (1993) Dynamic mutation in Dutch Huntington's disease patients: increased paternal repeat instability extending to within the normal size range. J Med Genet 30:996–1002

Duyao M, Ambrose C, Myers R, Noveletto A, Persichetti F, Frontali M, Dolstein S, et al (1993) Trinucleotide repeat length instability and age of onset in Huntington disease. Nat Genet 4:387–392

Goldberg YP, Andrew SE, Clarke LA, Hayden MR (1993a) A PCR method for accurate assessment of trinucleotide repeat expansion in Huntington disease. Hum Mol Genet 2:635–636

Goldberg YP, Kremer B, Andrew SE, Theilmann J, Graham

RK, Squitieri F, Telenius H, et al (1993b) Molecular analysis of new mutations causing Huntington disease: intermediate alleles and sex of origin effects. Nat Genet 5:174–179

Huntington's Disease Collaborative Research Group, The (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 72:971–983

Kremer B, Goldberg P, Andrew SE, Theilmann J, Telenius H, Zeisler J, Squitieri F, et al (1994) Worldwide study of the Huntington's disease mutation: the sensitivity and specificity of repeated CAG sequences. N Engl J Med 330:1401–1406

Kremer B, Squitieri F, Telenius H, Andrew SE, Theilmann J, Spence N, Goldberg YP, et al (1993) Molecular analysis of late onset Huntington's disease. J Med Genet 30:991–995

Kunkel LM, Smith KD, Boyer SH, Borgaonkar DS, Wachtel SS, Miller OJ, Breg WR, et al (1977) Analysis of human Y chromosome specific reiterated DNA in chromosome variants. Proc Natl Acad Sci USA 74:1245–1249

Myers RH, MacDonald ME, Koroshetz WJ, Duyao MP, Ambrose CM, Taylor SAM, Barnes G, et al (1993) *De novo* expansion of a (CAG)n repeat in sporadic Huntington's disease. Nat Genet 5:168–173

Novelletto A, Persichetti F, Sabbadini G, Mandich P, Bellone P, Bellone E, Ajmar F, et al (1994) Analysis of the trinucleotide repeat expansion in Italian families affected with Huntington disease. Hum Mol Genet 3:93–98

Snell R, MacMillan JC, Cheadle JP, Fenton I, Lazarou LP, Davies P, MacDonald ME, et al (1993) Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. Nat Genet 4:393–397

Squitieri F, Andrew SE, Goldberg YP, Kremer B, Spence N, Zeisler J, Nichol K, et al (1994) DNA haplotype analysis of Huntington disease reveals clues to the origins and mechanisms of CAG expansion and reasons for geographic variations of prevalence. Hum Mol Genet 3:2103–2114

Stine OC, Pleasant N, Franz ML, Abbott MH, Folstein SE, Ross CA (1993) Correlation between the onset age of Huntington's disease and length of the trinucleotide repeat in IT-15. Hum Mol Genet 2:1547–1549

Telenius H, Almqvist E, Kremer B, Spence N, Squitieri F, Nichol K, Grandell U (1995) Somatic mosaicism in sperm is associated with intergenerational (CAG)n changes in Huntington disease. Hum Mol Genet 4:189–195

Telenius H, Kremer HPH, Goldberg YP, Squitieri F, Theilmann J, Zeisler J, Adam S, et al (1994) Somatic and gonadal mosaicism in the Huntington disease gene CAG repeat in brain and sperm. Nat Genet 6:409–414

Telenius H, Kremer HPH, Theilmann J, Andrew SE, Almqvist E, Anvret M, Greenberg C, et al (1993) Molecular analysis of juvenile Huntington disease: the major influence of paternal descent on (CAG)n repeat length. Hum Mol Genet 2:1535–1545

Trottier Y, Biancalana, V, Mandel J-L (1994) Instability of CAG repeats in Huntington's disease: relation to parental transmission and age of onset. J Med Genet 31:377–382

Tukey JW (1977) Exploratory data analysis. Addison Wesley, Reading, MA

Zühlke C, Reiss O, Bockel B, Lange H, Thies U (1993) Mitotic stability and meiotic variability of the (CAG)n repeat in the Huntington disease gene. Hum Mol Genet 2:2063–2067